

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 13:56:00 ; Search time 26 Seconds
(without alignments)
576.131 Million cell updates/sec

Title: US-09-818-143-22

Perfect score: 539

Sequence: 1 MEALNVNFPKSRFPEDAG.....RLKVGPRQASQRLKETGLC 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubaa/FCI_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	539	100.0	99	9 US-09-818-143-22	Sequence 22, Appl
2	539	100.0	99	10 US-09-974-298-149	Sequence 149, Appl
3	79.5	14.7	322	9 US-09-789-836-6	Sequence 6, Appl
4	79.5	14.7	322	12 US-09-789-831-6	Sequence 6, Appl
5	79.5	14.7	345	9 US-09-789-836-5	Sequence 5, Appl
6	79.5	14.7	345	12 US-09-789-831-5	Sequence 5, Appl
7	79.5	14.7	345	12 US-10-251-661-10	Sequence 10, Appl
8	73.5	13.6	345	15 US-10-060-036-183	Sequence 183, Appl
9	70.5	13.1	419	16 US-10-138-098-25	Sequence 25, Appl
10	69.5	12.9	419	16 US-10-138-098-27	Sequence 27, Appl
11	69	12.8	323	15 US-10-205-823-62	Sequence 62, Appl
12	69	12.8	323	15 US-10-177-293-53	Sequence 53, Appl
13	69	12.8	4123	15 US-10-213-509-5	Sequence 5, Appl
14	67.5	12.5	627	12 US-10-349-836-6	Sequence 6, Appl
15	67	12.4	597	9 US-09-853-386-93	Sequence 93, Appl

Sequence 115, Appl
Sequence 6, Appl
Sequence 1316, Appl
Sequence 863, Appl
Sequence 12, Appl
Sequence 58, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 102, Appl
Sequence 61, Appl
Sequence 61, Appl
Sequence 13173, A
Sequence 20, Appl
Sequence 36, Appl
Sequence 14, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 827, Appl
Sequence 30, Appl
Sequence 5611, Appl
Sequence 103, Appl
Sequence 94, Appl
Sequence 206, Appl
Sequence 12, Appl
Sequence 20, Appl
Sequence 18, Appl
Sequence 4323, Appl
Sequence 1546, Appl
Sequence 198, Appl
Sequence 198, Appl

494 14 US-10-108-605-115
627 10 US-09-892-985-6
375 10 US-09-764-864-1316
378 10 US-09-764-864-863
559 10 US-09-877-633-12
230 9 US-09-880-192-58
230 12 US-10-427-348-58
230 12 US-10-260-715-8
1259 15 US-10-260-715-8
144 9 US-09-789-831-7
351 9 US-09-350-874-61
351 15 US-10-106-989-61
594 15 US-10-156-761-13173
668 9 US-09-730-989-20
880 11 US-09-893-519A-36
2783 10 US-09-816-669A-14
147 9 US-09-789-836-7
147 12 US-09-789-831-7
184 10 US-09-764-864-827
398 11 US-09-291-417-30
398 11 US-09-291-417-30
441 10 US-09-738-626-5611
591 11 US-09-291-417-103
948 9 US-09-888-615-94
492 9 US-09-764-898-206
616 12 US-10-251-661-12
668 12 US-10-240-154-20
704 12 US-10-240-154-18
103 15 US-10-106-698-4323
163 10 US-09-925-300-1546
419 12 US-10-200-562-188
419 12 US-10-237-551-198

ALIGNMENTS

RESULT 1

US-09-818-143-22
Sequence 22, Application US/09818143
Patent No. US20020019000A1
GENERAL INFORMATION:
APPLICANT: Walker, Michael G.
APPLICANT: Volkmut, Wayne
APPLICANT: Klingler, Tod M.
TITLE OF INVENTION: POLYNUCLEOTIDES COEXPRESSED WITH MATRIX-REMODELING GENES
FILE REFERENCE: PB-0004 CIP
CURRENT APPLICATION NUMBER: US/09/818.143
CURRENT FILING DATE: 2001-03-26
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PERL Program
SEQ ID NO 22
LENGTH: 99
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE: -
OTHER INFORMATION: 1556751CD1
US-09-818-143-22

Query Match 100.0%; Score 539; DB 9; Length 99;
Best Local Similarity 100.0%; Pred. No. 4.4e-48;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MEALNVNFPKSRFPEDAGKESGGGFCVPAARPTMTGTPSCSPGLNFSPOKEN 60

Db 1 MEALNVNFPKSRFPEDAGKESGGGFCVPAARPTMTGTPSCSPGLNFSPOKEN 60

Oy 61 RACACQWAGPAPKPMCMVRLKVGPRQASQRLKETGLC 99

Db 61 RACACQWAGPAPKPMCMVRLKVGPRQASQRLKETGLC 99

RESULT 2

US-09-974-298-149

; Sequence 149, Application US/09974298

```

; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huel-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1556751C01
; US-09-974-298-149

Query Match      100.0%; Score 539; DB 10; Length 99;
Best Local Similarity 100.0%; Pred. No. 4.4e-48;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEALANVFPKSFREDAGKESGQGFVPAARPOTMTGPGSCSPGLQNFSPORKEN 60
Db 1 MEALANVFPKSFREDAGKESGQGFVPAARPOTMTGPGSCSPGLQNFSPORKEN 60
QY 61 RACACWQAGPAPKPMCVRLKVGPRPOASORKLKETGLC 99
Db 61 RACACWQAGPAPKPMCVRLKVGPRPOASORKLKETGLC 99

RESULT 3
US-09-789-836-6
; Sequence 6, Application US/09789836
; Patent No. US20020082204A1
; GENERAL INFORMATION:
; APPLICANT: BRIGHAM, KENNETH L.
; APPLICANT: STECENKO, ARLENE A.
; APPLICANT: SEALY, LINDA
; TITLE OF INVENTION: TREATMENT OF INFLAMMATION WITH P20
; FILE REFERENCE: N-6977
; CURRENT APPLICATION NUMBER: US/09/789,836
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/183,584
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-789-836-6

Query Match      14.7%; Score 79.5; DB 9; Length 322;
Best Local Similarity 26.2%; Pred. No. 2.5;
Matches 32; Conservative 10; Mismatches 39; Indels 41; Gaps 5;

QY 10 PRKSFREDAGKESGQGFVPAARPOTMTGPGSCSPGLQNFSP 55
Db 158 PADCKRKEEAGAGGGAGMAGFPYALRAYLGYQAVPSGSGSLSTSSSSPPGTP--SP 215
QY 56 QRKENRACACWQAGPAP-----CVPAAARPOTMTGPGSCSPGLQNFSP 55
Db 216 ADAPKAPTACYAGAGPAPSOVKSKAKKTVDKHSDEYKIRERNIAVR--KSRDKAKMRN 273
QY 93 LK 94
Db 274 LE 275

US-09-789-831-6
; Sequence 6, Application US/09789831
; Patent No. US20030166586A1
; GENERAL INFORMATION:
; APPLICANT: SEALY, LINDA
; TITLE OF INVENTION: C/EBP-BETA ISOFORMS AND METHODS OF USE IN CELL REGULATION
; FILE REFERENCE: N-6978
; CURRENT APPLICATION NUMBER: US/09/789,831
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/183,532
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-789-831-6

Query Match      14.7%; Score 79.5; DB 12; Length 322;
Best Local Similarity 26.2%; Pred. No. 2.5;
Matches 32; Conservative 10; Mismatches 39; Indels 41; Gaps 5;

QY 10 PRKSFREDAGKESGQGFVPAARPOTMTGPGSCSPGLQNFSP 55
Db 158 PADCKRKEEAGAGGGAGMAGFPYALRAYLGYQAVPSGSGSLSTSSSSPPGTP--SP 215
QY 56 QRKENRACACWQAGPAP-----CVPAAARPOTMTGPGSCSPGLQNFSP 55
Db 216 ADAPKAPTACYAGAGPAPSOVKSKAKKTVDKHSDEYKIRERNIAVR--KSRDKAKMRN 273
QY 93 LK 94
Db 274 LE 275

US-09-789-836-5
; Sequence 5, Application US/09789836
; Patent No. US20020082204A1
; GENERAL INFORMATION:
; APPLICANT: BRIGHAM, KENNETH L.
; APPLICANT: STECENKO, ARLENE A.
; APPLICANT: SEALY, LINDA
; TITLE OF INVENTION: TREATMENT OF INFLAMMATION WITH P20
; FILE REFERENCE: N-6977
; CURRENT APPLICATION NUMBER: US/09/789,836
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/183,584
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-789-836-5

Query Match      14.7%; Score 79.5; DB 9; Length 345;
Best Local Similarity 26.2%; Pred. No. 2.7;
Matches 32; Conservative 10; Mismatches 39; Indels 41; Gaps 5;

QY 10 PRKSFREDAGKESGQGFVPAARPOTMTGPGSCSPGLQNFSP 55
Db 181 PADCKRKEEAGAGGGAGMAGFPYALRAYLGYQAVPSGSGSLSTSSSSPPGTP--SP 238
QY 56 QRKENRACACWQAGPAP-----CVPAAARPOTMTGPGSCSPGLQNFSP 55
Db 239 ADAPKAPTACYAGAGPAPSOVKSKAKKTVDKHSDEYKIRERNIAVR--KSRDKAKMRN 296
QY 93 LK 94
Db 274 LE 275

RESULT 4
US-09-789-836-6
; Sequence 6, Application US/09789836
; Patent No. US20020082204A1
; GENERAL INFORMATION:
; APPLICANT: BRIGHAM, KENNETH L.
; APPLICANT: STECENKO, ARLENE A.
; APPLICANT: SEALY, LINDA
; TITLE OF INVENTION: TREATMENT OF INFLAMMATION WITH P20
; FILE REFERENCE: N-6977
; CURRENT APPLICATION NUMBER: US/09/789,836
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/183,584
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-789-836-6

Query Match      14.7%; Score 79.5; DB 9; Length 322;
Best Local Similarity 26.2%; Pred. No. 2.5;
Matches 32; Conservative 10; Mismatches 39; Indels 41; Gaps 5;

QY 10 PRKSFREDAGKESGQGFVPAARPOTMTGPGSCSPGLQNFSP 55
Db 158 PADCKRKEEAGAGGGAGMAGFPYALRAYLGYQAVPSGSGSLSTSSSSPPGTP--SP 215
QY 56 QRKENRACACWQAGPAP-----CVPAAARPOTMTGPGSCSPGLQNFSP 55
Db 216 ADAPKAPTACYAGAGPAPSOVKSKAKKTVDKHSDEYKIRERNIAVR--KSRDKAKMRN 273
QY 93 LK 94
Db 274 LE 275
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 13:55:22 ; Search time 30 Seconds
(without alignments)
139.626 Million cell updates/sec

Title: US-09-818-143-22

Perfect score: 539

Sequence: 1 MEALANVFPKSRPEDAG.....RLKVGPRQASQRLKETGLC 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	77.5	14.4	300	4	US-09-252-991A-28367
2	76.5	14.2	253	4	US-09-252-991A-32747
3	74.5	13.8	138	4	US-09-252-991A-23683
4	74.5	13.8	298	4	US-09-252-991A-31845
5	74	13.7	176	4	US-09-252-991A-19223
6	73	13.5	683	4	US-09-252-991A-23859
7	72.5	13.5	181	4	US-09-252-991A-23483
8	72	13.4	412	4	US-09-252-991A-19536
9	72	13.4	1034	4	US-09-252-991A-28921
10	71.5	13.3	177	4	US-09-252-991A-20487
11	71	13.2	136	4	US-09-252-991A-26248
12	71	13.2	160	4	US-09-252-991A-17039
13	71	13.2	209	4	US-09-252-991A-23583
14	70.5	13.1	303	4	US-09-252-991A-19799
15	70.5	13.1	424	4	US-09-252-991A-32212
16	70	13.0	100	4	US-09-252-991A-31190
17	70	13.0	195	4	US-09-252-991A-28568
18	70	13.0	804	4	US-09-252-991A-21676
19	69.5	12.9	149	4	US-09-252-991A-32070
20	69.5	12.9	160	4	US-09-252-991A-22982
21	69.5	12.9	507	1	PCT-US95-0748-23
22	69.5	12.9	507	5	PCT-US95-05758-23
23	69	12.8	116	4	US-09-252-991A-32519
24	69	12.8	151	4	US-09-252-991A-25555
25	69	12.8	249	4	US-09-252-991A-22724
26	69	12.8	239	4	US-09-252-991A-17588
27	69	12.8	435	4	US-09-252-991A-18689

ALIGNMENTS

RESULT 1

US-09-252-991A-28367
; Sequence 28367, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252.991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 28367

LENGTH: 300

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-28367

Query Match 14.4%; Score 77.5; DB 4; Length 300;

Best Local Similarity 30.1%; Pred. No. 0.32; Indels 15; Gaps 5;

Matches 28; Conservative 10; Mismatches 40; Indels 15; Gaps 5;

QY 9 FPKSRPEDAG-----KESGSGGFCVPAARPTMTGTPSCSPGLQNFSPORKENRA 62

DB 129 FPOGKRPAAGLQPGAGLRQGVHRG-ARPSARPHGLRTGSLVLPGRPR--GTDREGRV 185

QY 63 CACQWAGPAPKPMCVRLKVGPRQASQRLKE 95

DB 186 PA-WQ-----PRPPVAARLPARQVRVGRLAQ 212

RESULT 2

US-09-252-991A-32747

; Sequence 32747, Application US/09252991A

; Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252.991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

Sequence 22557, A
Sequence 31872, A
Sequence 28124, A
Sequence 30927, A
Sequence 32059, A
Sequence 29281, A
Sequence 27781, A
Sequence 31279, A
Sequence 31016, A
Sequence 24759, A
Sequence 6, Appl1
Sequence 28968, A
Sequence 31890, A
Sequence 22641, A
Sequence 20033, A
Sequence 18366, A
Sequence 17070, A

4 US-09-252-991A-22557
4 US-09-252-991A-31872
4 US-09-252-991A-28124
4 US-09-252-991A-30927
4 US-09-252-991A-32059
4 US-09-252-991A-29281
4 US-09-252-991A-27781
4 US-09-252-991A-31279
4 US-09-252-991A-31016
4 US-09-252-991A-24759
4 US-08-487-536-6
4 US-08-660-451A-6
4 US-09-252-991A-28968
4 US-09-252-991A-31890
4 US-09-252-991A-22641
4 US-09-252-991A-20033
4 US-09-252-991A-18366
4 US-09-252-991A-17070

28 68.5 12.7 191
29 68.5 12.7 215
30 68.5 12.7 251
31 68.5 12.7 295
32 67.5 12.5 151
33 67.5 12.5 247
34 67.5 12.5 250
35 67.5 12.5 268
36 67.5 12.5 362
37 67.5 12.5 488
38 67.5 12.5 627
39 67.5 12.5 627
40 67 12.4 133
41 67 12.4 151
42 67 12.4 208
43 67 12.4 420
44 67 12.4 466
45 67 12.4 539

PD 20-APR-2000.

XX PF 06-OCT-1999; 99WO-US23315.
 XX PR 09-OCT-1998; 98US-0169289.
 XX PA (INCY-) INCYTE PHARM INC.
 XX PI Walker MG, Volkmuth W, Klingler TM;
 XX DR WPI: 2000-317934/27.
 XX DR N-PSDB; AAZ52354.
 XX PT Protein co-expressed with matrix-remodeling proteins, useful in the
 PT diagnosis and treatment of cancer, cardiomyopathy, arthritis,
 PT angiogenesis, diabetic necrosis, atherosclerosis, fibrosis, and
 PT ulceration.
 XX PS Claim 1: Page 53-54; 55pp; English.
 XX CC The present sequence is PSEQ protein encoded by NSEQ gene that is
 C co-expressed with one or more known matrix-remodeling genes in a number
 CC of biological samples using an expression vector. The gene, protein, and
 CC antibody sequences can be used in the diagnosis, and treatment or
 CC prevention of a disease associated with its altered expression. The
 CC diseases that can be treated are matrix-remodeling diseases, including
 CC cancer, cardiomyopathy, arthritis, angiogenesis, diabetic necrosis,
 CC atherosclerosis, fibrosis, and ulceration.
 XX SQ Sequence 99 AA;
 Query Match 100.0%; Score 539; DB 21; Length 99;
 Best Local Similarity 100.0%; Pred. No. 1.7e-51;
 Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEALANVFPKSRFPEDAGKESGQGFVCPAARPTMTGTPSCSSPGLQNFSPQRKEN 60
 Db 1 MEALANVFPKSRFPEDAGKESGQGFVCPAARPTMTGTPSCSSPGLQNFSPQRKEN 60
 QY 61 RACACWQAGPAPKPMCVRLKVGPRQASQKRLKETGLC 99
 Db 61 RACACWQAGPAPKPMCVRLKVGPRQASQKRLKETGLC 99
 RESULT 2
 AA014785
 ID AA014785 standard; Protein; 99 AA.
 XX AC AA014785;
 X DT 11-JUL-2002 (first entry)
 XX DE Human matrix-remodeling-associated protein 2.
 XX KW Human; matrix-remodeling gene; extracellular matrix;
 KW matrix-remodeling-associated nucleotide; screening;
 KW matrix remodelling-associated disease; angiogenesis; arthritis;
 KW atherosclerosis; cancer; cardiomyopathy; diabetic necrosis; fibrosis;
 KW ulceration.
 XX OS Homo sapiens.
 XX PN US2002019000-A1.
 XX PD 14-FEB-2002.
 XX PF 26-MAR-2001; 2001US-0818143.
 XX PR 09-OCT-1998; 98US-0169289.
 XX PA (WALK/) WALKER M G.
 PA (VOLK/) VOLKMUTH W.
 PA (KLIN/) KLINGLER T M.
 XX PT

PI Walker MG, Volkmuth W, Klingler TM;
 XX DR WPI: 2002-338319/37.
 XX DR N-PSDB; AAL42450.
 XX PT New isolated polynucleotide coexpressed with matrix-remodeling genes,
 PT useful in diagnosis, prognosis, prevention and treatment of diseases
 PT associated with matrix-remodeling such as angiogenesis, arthritis and
 PT cancer.
 XX PS Examples; Fig 2; 63pp; English.
 XX CC The invention comprises human nucleotide sequences which are co-expressed
 CC with matrix-remodeling genes. Matrix-remodeling is associated with the
 CC construction, destruction and reorganisation of extracellular matrix
 CC components. The matrix-remodeling-associated nucleotides of the invention
 CC are useful for screening for and purifying ligands that specifically bind
 CC to the nucleotides of the invention. The matrix-remodeling-associated
 CC nucleotides of the invention are also useful in the diagnosis, prognosis,
 CC prevention, treatment and evaluation of therapies for diseases associated
 CC with matrix remodelling (e.g. angiogenesis, arthritis, atherosclerosis,
 CC cancer, cardiomyopathy, diabetic necrosis, fibrosis and ulceration). The
 CC present amino acid sequence is encoded by a human matrix-remodeling.
 CC associated nucleotide of the invention.
 XX SQ Sequence 99 AA;
 Query Match 100.0%; Score 539; DB 23; Length 99;
 Best Local Similarity 100.0%; Pred. No. 1.7e-51;
 Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEALANVFPKSRFPEDAGKESGQGFVCPAARPTMTGTPSCSSPGLQNFSPQRKEN 60
 Db 1 MEALANVFPKSRFPEDAGKESGQGFVCPAARPTMTGTPSCSSPGLQNFSPQRKEN 60
 QY 61 RACACWQAGPAPKPMCVRLKVGPRQASQKRLKETGLC 99
 Db 61 RACACWQAGPAPKPMCVRLKVGPRQASQKRLKETGLC 99
 RESULT 3
 ABUS7648
 ID ABUS7648 standard; Protein; 99 AA.
 XX AC ABUS7648;
 XX DT 09-APR-2003 (first entry)
 XX DE Differentially expressed breast cancer associated protein #35.
 XX KW Breast cancer; differential gene expression; BC-cDNA;
 KW breast cancer diagnosis; breast cancer monitoring;
 KW breast cancer treatment; breast cancer staging.
 XX OS Homo sapiens.
 XX PN US2002156263-A1.
 XX PD 24-OCT-2002.
 XX PF 04-OCT-2001; 2001US-0974298.
 XX PR 05-OCT-2000; 2000US-238331P.
 XX PA (CHEN/) CHEN H.
 XX PI Chen H;
 XX WPI: 2003-182653/18.
 XX New cDNAs, which are differentially expressed in (metastatic) breast
 PT cancer useful for diagnosing or staging, breast cancer, or for
 PT monitoring the treatment of breast cancer in an individual.

Result No.	Score	Query Match	Length	DB ID	Description
1	339	100.0	99	4 Q9UFS0	Q9UFS0 homo sapien
2	86.5	16.0	537	3 Q9U094	Q9U094 schizosacchi
3	83.5	13.5	506	5 Q8T433	Q8T433 drosophila
4	78	14.5	194	5 Q8SLZ2	Q8SLZ2 drosophila
5	76.5	14.5	504	5 Q8MYR8	Q8MYR8 drosophila
6	76.5	14.2	208	13 Q9J5Y92	Q9J5Y92 astyanax fa
7	75.5	14.0	1534	4 Q9VEH3	Q9VEH3 drosophila
8	75	13.9	712	4 Q8NCN4	Q8NCN4 homo sapien
9	74.5	13.8	368	4 Q96W84	Q96W84 homo sapien
10	74.5	13.6	345	4 Q96LH2	Q96LH2 homo sapien
11	73	13.5	386	10 Q85ZF1	Q85ZF1 oryza sativa
12	73	13.5	945	5 Q9NKT7	Q9NKT7 leishmania
13	73	13.5	2087	5 Q8MXL2	Q8MXL2 leishmania
14	72.5	13.5	323	4 Q8TD94	Q8TD94 homo sapien
15	72.5	13.5	504	5 P90730	P90730 ceratit c
16	72	13.4	2533	5 P90589	P90589 parametium

61 RACACWONAGPAPKNPMCVRILKVGRRPOASORKIKETGIC 99

```

RESULT 2
Q9U094 ID Q9U094 PRELIMINARY; PRT; 557 AA.
AC Q9U094;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical 64.6 kDa protein.
GN SPC830.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h;
A McDougall R.M., Rajandream M.A., Barrell B.G., Ransperger U., Pohl T.;
L Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL109830; CAB52878.1; -
DR GenDB: SPombe; SPC830.05C; -
KW Hypothetical protein.
SQ SEQUENCE 557 AA; 64627 MW; 884CE3D8559EA724 CRC64;

Query Match 16.0%; Score 86.5; DB 3; Length 557;
Best Local Similarity 30.4%; Pred. No. 0.12;
Matches 28; Conservative 18; Mismatches 33; Indels 13; Gaps 6;

QY 6 NVNPRKSRF---PEDAGKES-GSGGFCVPAARQPMW7--GPGSSP-GLQNF---- 53
Db 402 HTNYPDFSYNIPIYSGRETKQSHNQLSFSSTPPLSDNGFTYSTPHSLSNFNTCD 461

QY 54 SPQKRNACACWQAGAPKPMCVRLKVG 85
Db 462 SLSPSSNNLYGYSTL-LHPRNPICVRQIGR 492

RESULT 3
Q8T433 ID Q8T433 PRELIMINARY; PRT; 1196 AA.
AC Q8T433;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE AT19506p.
GN CG14322.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY089375; AAL90113.1; -
DR Flybase; FBgn0038532; CG14322.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 4.
DR SMART; SM00320; WD40; 4.
DR Repeat; WD repeat.
SQ SEQUENCE 1196 AA; 133061 MW; 9ADBA0269C8DEB47 CRC64;

Query Match 15.5%; Score 83.5; DB 5; Length 1196;
Best Local Similarity 30.1%; Pred. No. 0.62;
Matches 28; Conservative 16; Mismatches 32; Indels 17; Gaps 4;

QY 2 EALANVFPRK-SFRPEDAGKES---GSGGFCVPAARQPMW7GP-----SC 45
Db 481 QKLANLLSPKKTAFSTDETKASTLCPNQHGFIPTAQDSSPYTSCQKPPSSPPSQSP 540

QY 46 SSPGLQNFSPQR-KENRACACWQAGAPKPM 77
Db 541 SSATAQSSPMALQKQSCPAQSCSTPIQENPM 573

RESULT 4
Q8SZL2 ID Q8SZL2 PRELIMINARY; PRT; 504 AA.
AC Q8SZL2;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE RH16046p.
GN CYP6A21 OR CG10247.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL: AY070678; AAL48149.1; -
DR Flybase; FBgn0033981; Cyp6a21.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 504 AA; 57842 MW; 32DA87A519DF557A CRC64;

Query Match 14.5%; Score 78; DB 5; Length 504;
Best Local Similarity 30.5%; Pred. No. 1.1;
Matches 18; Conservative 11; Mismatches 26; Indels 4; Gaps 2;

QY 30 CYPARPQPMVTFGSCSPGLQNFSPQRKRNACACWQAGAPKPMCVRLKVG 88
Db 405 CGAMHRDEKLYANPNTFNP--DNFSPVRKERSVEMLPFGDGPRN--CIGMRFGQMQA 459

RESULT 5
Q8MY8 ID Q8MY8 PRELIMINARY; PRT; 504 AA.
AC Q8MY8;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE SD12535p.
GN CYP6A9 OR CG10246.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY089375; AAL90113.1; -
DR Flybase; FBgn0038532; CG14322.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 4.
DR SMART; SM00320; WD40; 4.
DR Repeat; WD repeat.
SQ SEQUENCE 1196 AA; 133061 MW; 9ADBA0269C8DEB47 CRC64;

Query Match 15.5%; Score 83.5; DB 5; Length 1196;
Best Local Similarity 30.1%; Pred. No. 0.62;
Matches 28; Conservative 16; Mismatches 32; Indels 17; Gaps 4;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 13:55:21; Search time 23 Seconds

(without alignments)
202.419 Million cell updates/sec

Title: US-09-818-143-22

Perfect score: 539

Sequence: 1 MEALNVNPRKSRPEDAG.....RLKVGPRQASQRLKETGLC 99

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79.5	14.7	345	1	CEBB_HUMAN
2	79.5	14.7	998	1	RRPO_BOOLV
3	78.5	14.6	998	1	RRPO_FHV
4	78	14.5	504	1	C6A9_DROME
5	78	14.5	504	1	C6A1_DROME
6	78	14.5	825	1	SES_RAT
7	77	14.3	348	1	CEBB_BOVIN
8	75.5	14.0	998	1	RRPO_BBV
9	75	13.9	305	1	GSH2_MOUSE
10	74	13.7	119	1	VTU3_DROME
11	72	13.4	137	1	VIT_STRPU
12	72	13.4	740	1	POLB_MAIZE
13	70	13.0	2003	1	NTC4_HUMAN
14	69.5	12.9	506	1	C6A2_DROME
15	69	12.8	323	1	C6A8_HUMAN
16	68	12.6	569	1	SILF_MOUSE
17	68	12.6	433	1	FUT4_RAT
18	68	12.6	3726	1	ABF1_MOUSE
19	67.5	12.5	627	1	ACHA_HUMAN
20	67	12.4	597	1	NR41_RAT
21	67	12.4	1235	1	IRS1_RAT
22	66.5	12.3	95	1	Y95_ADE02
23	66.5	12.3	506	1	C6A8_DROME
24	66.5	12.3	1355	1	XDH_MOUSE
25	66	12.2	507	1	CGA1_DROME
26	65.5	12.2	1786	1	UVRA_CHLTR
27	65	12.1	172	1	CX42_RAT
28	65	12.1	559	1	DACA_HUMAN
29	65	12.0	606	1	E75_METEN
30	64.5	12.0	457	1	Z185_HUMAN
31	64	11.9	907	1	GP3_HUMAN
32	64	11.9	3703	1	ABF1_HUMAN
33	63.5	11.8	591	1	PAK4_HUMAN

34	63.5	11.8	813	1	YTOJ_CAEBL	Q19673 caenorhabdi
35	63	11.7	222	1	HXC5_HUMAN	Q00444 homo sapien
36	63	11.7	222	1	HXC5_MOUSE	P32043 mus musculus
37	63	11.7	387	1	GT4R_HUMAN	Q9nr83 homo sapien
38	63	11.7	390	1	VE2_CRPVK	P03121 cottontail
39	63	11.7	616	1	VGF_HUMAN	O15240 homo sapien
40	63	11.7	670	1	SYN1_MOUSE	O88935 mus musculus
41	63	11.7	704	1	SYN1_RAT	P09951 rattus norv
42	62.5	11.6	406	1	PLG2_HUMAN	Q9brq0 homo sapien
43	62.5	11.6	479	1	KLFA_MOUSE	O89091 mus musculus
44	62.5	11.6	598	1	NR41_CANFA	P51666 canis famil
45	62.5	11.6	686	1	DLA4_MOUSE	Q9J171 mus musculus

ALIGNMENTS

RESULT 1
CEBB_HUMAN
ID CEBB_HUMAN STANDARD; PRT: 345 AA.
AC P17676; Q9H425;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE CCAAT/enhancer binding protein beta (C/EBP beta) (Nuclear factor NF-IL6) (Transcription factor 5).
DE NF-IL6 (Transcription factor 5).
GN CEBPB OR TCF5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCB:TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=90269225; PubMed=2112087;
RA Akira S., Ishihara H., Sugita T., Tanabe O., Kinoshita S., Nishio Y., Nakajima T., Hirano T., Kishimoto T.
RT "A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family."
RL EMBL J. 9:1897-1906(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavridis G., Almeida J.P., Babbage A.K., Baggeley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clegg S., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R., Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Lehesvallo M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McConachie L.J., McIlroy K., McMurray A., Milne S.A., Mistry D., Moore N.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H., Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S., Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.,
RT "The DNA sequence and comparative analysis of human chromosome 20."
RL Nature 414:865-871(2001).
CC -!- FUNCTION: IMPORTANT TRANSCRIPTIONAL ACTIVATOR IN THE REGULATION OF GENES INVOLVED IN IMMUNE AND INFLAMMATORY RESPONSES. SPECIFICALLY BINDS TO AN IL-1 RESPONSE ELEMENT IN THE IL-6 GENE. NF-IL6 ALSO

```

Viruses; ssRNA positive-strand viruses, no DNA stage; Nodaviridae;
CC Alphandonavirus.
CC NCBI_TaxID=12286;
CC [1]
CC RN RP SEQUENCE FROM N.A.
CC RA MEDLINE=21351039; PubMed=11457991;
CC RX Johnson K.N., Johnson K.L., Dasgupta R., Ball L.A.;
CC RT "Comparisons among the larger genome segments of six nodaviruses and
CC their encoded RNA replicases";
CC RL J. Gen. Virol. 82:1855-1868(2001).
CC CC -1- FUNCTION: Replicates the viral genome which is composed of two RNA
CC segments, RNA1 and RNA2.
CC CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC CC -1- SIMILARITY: BELONGS TO THE NODAVIRUSES RNA POLYMERASE FAMILY.
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
CC CC EMBL: AF329080; AAK15751.1; -
CC DR InterPro: IPR007095; RNA_pol_DS_PS.
CC KW Transferase; RNA-directed RNA polymerase.
CC SQ SEQUENCE 998 AA; 111203 MW; 7FAFFB4FC7329253 CRC64;
CC
CC Query Match 14.7%; Score 79.5; DB 1; Length 998;
CC Best Local Similarity 28.2%; Pred. No. 2.1;
CC Matches 33; Conservative 12; Mismatches 31; Indels 41; Gaps 7;
CC
CC QY 7 VNPRKS---PRPEDA-GKSSS-----QGFCV-----PAARPOTM-VTG 42
CC Db 628 INCPAAKRFEPGPGMGVKSSTPTTHNTQYNACVEYTAALKFFYPDANPETLFLSLG 687
CC
CC QY 43 PSCSSGGLQNSPQRKENACACQWAGPAPKPMCVRLKVGPRQASQRLKGTGLC 99
CC Db 688 PKCGDDGLARATIKNTINRAAKCYG-----LELKV-----EKNPEVGLC 727
CC
CC RESULT 3
CC RRPO_FHV
CC ID RRPO_FHV STANDARD; PRT; 998 AA.
CC AC Q66929;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE RNA-directed RNA polymerase (EC 2.7.7.48) (RdRp) (RNA replicase)
CC DE (Protein A).
CC OS Flock house virus (FHV).
CC OS Viruses; ssRNA positive-strand viruses, no DNA stage; Nodaviridae;
CC CC Alphandonavirus.
CC CC NCBI_TaxID=12287;
CC CC [1]
CC RN RP SEQUENCE FROM N.A.
CC RA Dasgupta R.;
CC RT "Near identity in the polymerase gene of two serologically distinct
CC nodaviruses.";
CC RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
CC CC -1- FUNCTION: Replicates the viral genome which is composed of two RNA
CC segments, RNA1 and RNA2.
CC CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC CC -1- SIMILARITY: BELONGS TO THE NODAVIRUSES RNA POLYMERASE FAMILY.
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2003, 13:55:22 ; Search time 40 Seconds
(without alignments)
238.017 Million cell updates/sec

Title: US-09-818-143-22

Perfect score: 539

Sequence: 1 MEALNVNFPKSRPEDAG.....RLKVGPRQASOKLKETGLC 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:.*
1: piri:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	539	100.0	99	2	T17268
2	86.5	16.0	557	2	T41631
3	79.5	14.7	345	2	S12788
4	78.5	14.6	598	2	S41397
5	78	14.5	825	2	Jc4163
6	75.5	14.0	998	1	QBBB81
7	75	13.9	305	2	I57039
8	74	13.7	104	2	A22706
9	72	13.4	137	2	S00696
10	72	13.4	740	2	T03925
11	72	13.4	2333	2	T28675
12	72	13.4	2533	2	T28674
13	71.5	13.3	401	2	T02380
14	70.5	13.1	3942	2	T42730
15	69.5	12.9	507	1	A47198
16	69.5	12.9	511	2	S24345
17	68	12.6	751	2	F87789
18	67.5	12.5	98	2	F83404
19	67.5	12.5	627	2	Jc4021
20	67	12.4	563	2	J00623
21	67	12.4	1335	1	S16948
22	66.5	12.3	75	2	S05589
23	66.5	12.3	95	2	A03858
24	66.5	12.3	294	2	S13141
25	66.5	12.3	494	2	Jc5320
26	66.5	12.3	906	2	T28034
27	66.5	12.3	1335	1	X0MSDH
28	66	12.2	108	2	S37139
29	66	12.2	457	2	E88456

30	66	12.2	670	2	T49510	fibroin-3 related
31	65	12.2	1101	2	T20881	hypothetical prote
32	65.5	12.2	372	2	F86189	hypothetical prote
33	65.5	12.2	1786	1	H1527	probable excluducl
34	65	12.1	845	2	T17291	hypothetical prote
35	64.5	12.0	3938	2	T42761	Bassoon protein -
36	64	11.9	402	2	I46053	connexin44 - bovin
37	64	11.9	2783	1	A41948	alpha-fetoprotein
38	63.5	11.8	360	2	S75453	hypothetical prote
39	63.5	11.8	596	2	F75302	ATP-dependent DNA
40	63.5	11.8	649	2	S67787	hypothetical prote
41	63.5	11.8	736	2	T41259	hypothetical prote
42	63.5	11.8	813	2	T21132	hypothetical prote
43	63	11.7	222	2	S20029	homeotic protein H
44	63	11.7	303	2	T28999	hypothetical prote
45	63	11.7	390	1	WZLRR	E2 protein - cotto

ALIGNMENTS

RESULT 1

T17268

hypothetical protein DKFZp434N161.1 - human

C:Species: Homo sapiens (man)

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T17268

R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, September 1999

A:Reference number: Z18723

A:Accession: T17268

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-99 <POU>

A:Cross-references: EMBL:AL117489

A:Experimental source: adult testis; clone DKFZp434N161

C:Genetics:

A:Note: DKFZp434N161.1

Query Match 100.0% Score 539; DB 2; Length 99;

Best Local Similarity 100.0%; Pred. No. 2.6e-47;

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEALNVNFPKSRPEDAGKESGGGFCVPAARQTNMTGPGSCSSPGLQNFSPORKEN 60

Db 1 MEALNVNFPKSRPEDAGKESGGGFCVPAARQTNMTGPGSCSSPGLQNFSPORKEN 60

Qy 61 RACACWQAGPAPKPNMPCVRLKVRPQASOKLKETGLC 99

Db 61 RACACWQAGPAPKPNMPCVRLKVRPQASOKLKETGLC 99

RESULT 2

T41631

hypothetical protein SPCC830.05c - fission yeast (Schizosaccharomyces pombe).

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T41631

R:McDougal, R.M.; Rajandream, M.A.; Barrell, B.G.; Ramsperger, U.; Pohl, T.

submitted to the EMBL Data Library, August 1999

A:Reference number: Z22005

A:Accession: T41631

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-557 <MCD>

A:Cross-references: EMBL:AL109850; PIDN:CA852878.1; GSPDB:GN00068; SPDB:SPCC830.05c

A:Experimental source: strain 972h-; cosmid c830

C:Genetics:

A:Gene: SPDB:SPCC830.05c

A:Map position: 3

A:Introns: 56/3

Query Match 16.0% Score 86.5; DB 2; Length 557;

